UNCLASSIFIED

ODP 80-1336

ORD-1190-80
7 OCT 1980

MEMORANDUM FOR: Director of Data Processing

FROM : Director of Research and Development

SUBJECT : Distribution of ORD Technical Papers

1. Regarding previous briefings to members of your staff on several of ORD's FY-81 programs, we have since had several technical papers written covering many of the items discussed. We thought you would be interested in seeing this paper.

2. Attached you will find the technical paper generated by ORD which relates to your area of interest.

- Mathematical Programming: A Set of Tools Which Have Great Utility and Potential for the Intelligence Community. (U) (Author - [redacted]

STATINTL

3. Please feel free to contact the author directly for any additional information and/or comments.

STATINTL

[redacted]

Attachment:
As Stated

UNCLASSIFIED

MATHEMATICAL PROGRAMMING:

A SET OF TOOLS WHICH HAVE GREAT UTILITY AND

POTENTIAL FOR THE INTELLIGENCE COMMUNITY

AMR/ORD/DDS&T

DERIVATIVE CL BY 482859
☐ DECL ☒ REVW ON AUG 2000
DERIVED FROM C9e(1.5)

ABSTRACT

This paper illustrates some of the successful applications of mathematical programming techniques to problems of the military, industry, and Intelligence Community. Since linear programming has played a significant role in analyzing problems of the petroleum industry, a close examination of some of the types of problems of this industry is included in this report. (A simplified linear programming model of a refinery is developed in Appendix A.) Some of ORD's projects which utilize mathematical programming techniques are illustrated as well as the approaches which ORD has taken in expanding the use of these tools throughout the Agency. Appendix B provides a brief mathematical development of linear programming and some geometrical interpretations.

I. Mathematical Programming: An Historical Perspective

Mathematical programming is a term that was developed around 1950. It is now a generic term which encompasses linear programming, integer programming, dynamic programming, nonlinear programming, programming under uncertainty, and network flow problems. However, a class of mathematical programming problems can be traced back to the 17th century. The classical problems of determining the points of maxima and minima for constrained and unconstrained functions are one type of mathematical programming problems. A mathematical programming problem can be defined as maximizing (or minimizing) a multivariate objective function whose solution must satisfy a system of constraining functions. Mathematically, the problem can be written as follows:

$$\text{MAXIMIZE } f(X_1, X_2, \ldots, X_n)$$
$$\text{subject to } g_i(X_1, X_2, \ldots, X_n) \leq 0 \text{ for } i = 1, 2, \ldots, m.$$

The problems currently characterized as mathematical programming problems have the characteristic that the optimal solution cannot be obtained by a closed form procedure such as differential calculus. These types of problems are solved by well-defined computational algorithms. The algorithms are iterative in nature and usually yield an improved solution at each iteration. If the algorithm converges, then the final iteration will provide the optimal solution.

Mathematical programming techniques are general in the sense that they can be applied to models from many different environments.

Mathematical programming tools are one of the main analytical methodologies that are currently employed by operations research analysts. Many of the military operations research problems that were formulated in World War II could not be solved by proven algorithmic procedures, but were solved by heuristic or informal approaches. The planning tools of World War II ultimately led to the development of scientific programming procedures in the post-war period. Concurrent advances in the development of computers also greatly influenced these developments. Probably, the most significant development in the history of operations research was the development in 1947 by George B. Dantzig of the simplex algorithm for solving linear programming problems. The simplex algorithm provided the computational method for efficiently solving linear programming problems. Electronic digital computers quickly became the tools for the application of this approach in areas where hand computation was infeasible.

In the early 1950's, operations research analysts extended applications of mathematical programming techniques from the military problems to business and industrial sectors. By the early 1960's, almost all of the theoretical work in linear programming had been completed, and there were thousands of practical applications of this technique. During this decade, the emphasis was on developing the theoretical properties, computational procedures, and applications of the other mathematical programming formulations. The major advances in linear programming dealt with computational enhancements and more extensive applications.

The technological advances of the computer industry have had a profound influence on many disciplines, but they have affected none more dramatically than operations research. Without these capabilities, many mathematical programming techniques would be merely theoretical niceties. The advent of the computer has given rise to the development of computer-based modelling systems. The techniques for building, generating, solving, refining, and analyzing such models have undergone a steady evolutionary development as computer hardware progressed.

Evolutions in computers and computer-based modelling systems during the 1970's spawned new developments for modelling, solving, and efficiently implementing new solution procedures for network problems. From the early days of mathematical programming, practitioners had long recognized that problems with network structures represented one of the most significant classes of linear and integer programming problems. The classical transportation (distribution) problem of shipping goods from many warehouses to large numbers of destinations at minimum cost falls into the category of network problems. The potential industrial applications for efficiently solving large-scale transportation problems and other network formulations such as scheduling, resource allocation, production and inventory management were a prime motivation for the new developments. Before 1970, the consensus among mathematical programmers was that there were no significant algorithmic refinements which would greatly enhance the solution procedures for network problems. In retrospect, this attitude seems surprising since only minimal research had been undertaken to determine the computational strengths and weaknesses of alternate approaches. The development of effective algorithmic

techniques to exploit the specialized structure of network and network-related problems has increased the speed of solving these problems by two or three orders of magnitude over the standard linear programming approach. An immediate result of these breakthroughs is the savings of thousands of dollars in computer costs for organizations which utilize these types of models. In fact, it has recently been reported that over one-fourth of all computer time devoted to scientific computation is consumed by mathematical programming techniques.

A 1977 survey of over one hundred large companies in the United States revealed that 79% of the companies used linear and network programming models for analyzing their corporate problems. The following table shows the usage of other mathematical programming techniques by these companies:

Project Evaluation and Review Technique (PERT)/ Critical Path Method (CPM)	70%
Inventory Modelling	57%
Nonlinear Programming	36%
Heuristic Programming	34%
Dynamic Programming	27%
Mixed Integer and Integer Programming	2%

The following table shows usage of mathematical programming techniques by application:

Production Scheduling	70%
Inventory Modelling	70%
Capital Budgeting	56%

Transportation	51%
Plant Location	42%
Advertising and Sales Research	35%
Equipment Replacement	33%
Maintenance and Repair	28%
Packaging	9%

II. Applications of Mathematical Programming

The long succession of mathematical programming formulations of military problems began during World War II. Mathematical programming formulations and techniques have continued to be vital and effective tools of military strategists to this present day. As new areas have developed in mathematical programming, there have often been applications to military problems.

The following list provides a sampling of mathematical programming applications to military problems; it is not an all-inclusive representation:

- Arsenal exchange models
- Optimal missle trajectories
- Tactical air-to-air combat models
- Optimal allocation of missiles against area and point defenses
- Airlift and tanker routing problems
- Manpower planning models for promotion, rotation, attrition, etc.
- Scheduling of maintenance overhaul cycles
- Production sequencing of complex weapons systems
- Multiperiod allocation of funds for R&D projects
- Military interdiction models
- Logistics planning models
- Scheduling of flight training and survival courses

If the military applications of mathematical programming are excluded from consideration, there is a noticeable difference in the types of analytical approaches that are utilized by

government and industry. Two particular areas of mathematical programming, linear programming and network modelling, have gained widespread acceptance and are extensively used throughout industry. However, these techniques are employed occasionally, but not to their fullest capacity by governmental organizations. Two possible reasons for this may be that the goals which are to be optimized are not as easy to represent functionally as those of industry and that the diffusion of authority and responsibility in government makes it difficult to apply these techniques.

A list of applications of mathematical programming to problems of private industry would be enormous. The following brief list illustrates the variety of applications:

- Harvesting of timber lands
- Selecting portfolios of stocks, bonds and other securities
- Determining the locations and quantities of emergency services
- Scheduling of heating oil production, storage, and distribution to meet uncertain weather situations
- Analyzing competition and mergers within an industry
- Decreasing the amount of trim losses in paper mills
- Planning of menus (nutrition) for schools, institutions, and hospitals
- Optimally designing pressure vessels
- Determining lot sizes for production of automobiles and assigning production of machines
- Exploring, drilling, and planning production for crude oil
- Shipping food to and from warehouses
- Designing communication networks for optimal flow of messages

Since many intelligence problems form a natural pairing with the tools of mathematical programming, the CIA, in recent years, has extended its utilization of mathematical programming for analyzing many intelligence problems. ORD has been a leader in exploiting, refining, and transferring these techniques into operational use. Several approaches have been taken by ORD in expanding the application of mathematical programming tools. The methods of extending this methodology to other offices have been by funding contractual research, developing analytical models "in-house", providing consulting expertise to other offices, and contributing previously developed software. While progress has been impressive, a multitude of unsolved intelligence problems are amenable to resolution through mathematical programming formulation.

The following is a partial list of R&D projects which have been previously supported or have recently been undertaken by ORD to support intelligence analysis:

- o A model for the interdiction of sea shipping lanes
- o A model of the Soviet military logistics system
- o A model assigning nuclear warheads to Soviet missile systems
- o A model for estimating production streams of a Soviet refinery
- o A model for determining the flow from all oil fields to all refineries in the Soviet Union
- o Algorithms for estimating the technological coefficients of an input-output model of the Chinese economy
- o Algorithms for developing statistical estimates used in image processing

- Algorithms for developing statistical estimators used in secret writing analysis
- Models for determining directions of growth or contraction in sectors of an economy of a country

III. Linear Programming In The Petroleum Industry And Its Utility For Intelligence Applications

The major integrated oil companies, small oil companies, and countries such as the Soviet Union which perform planning of their oil production and consumption at the state level are all heavily dependent on linear programming to aid in decision making. The petroleum industry's use of linear optimization models deserves special attention for two important reasons. First, oil companies throughout the world have the best overall record of success in early and continued application of linear programming. Their experience amply demonstrates that it is practical and profitable to use mathematical models for planning purposes. Second, the oil companies, encouraged by their initial success, have pioneered the expansion of linear optimization methods to a wide variety of decision areas, and thus have demonstrated techniques for making this scientific approach workable in a competitive business environment.

The petroleum industry must have the capability of analyzing problems in the entire flow sequence which starts with pumping crude from the earth and ends with marketing of a product such as gasoline at service stations. Linear programming models have been formulated and tested to aid in decision making at every major point in this stream. Specifically, models have been developed to:

- o Schedule production from a series of underground oil reservoirs to maximize profit, subject to equipment capacity limitations and constraints imposed by physical pumping phenomena;

- o Determine the net profitability of exchanging a proprietary crude for another company's crude, given the configuration of refineries and the associated economics of processing the exchanged crude;

- o Calculate the incremental cost of increasing the amount of a product for a spot sale (for example, manufacturing a specified amount of jet fuel for a government contract), given the targeted quantities of other products to be manufactured;

- o Plan weekly minimum-cost schedules for refinery unit operations and product-blending, taking into account crude availabilities, throughput constraints on the refining units, performance characteristics of each product (such as octane rating), and the pre-established shipping requirements for the products;

- o Establish the return on investment of a proposed new refinery unit, to realize its full impact on existing units;

- o Route products from several refineries to a number of marketing areas along least-cost transportation paths, recognizing factors of differential costs of manufactured products at the separate refineries, relative shipping charges, and seasonal variations in customer demand;

- o Construct an annual plan to integrate the major decisions of the entire company.

The list of linear programming applications has not been exhausted for the petroleum industry, and there are many more that are as important as the ones presented. The various oil companies also utilize linear programs in different ways. Since each oil company has its peculiar characteristics, such as refinery location and age configuration, marketing districts, crude-oil reserves, and so forth, firms differ in their use of linear programming models. Some find it convenient to have one or two comprehensive models that

can be repeatedly employed to make several of the analyses described above. Other companies have constructed separate models for each specific purpose, with differing degrees of complexity and detail.

Certain patterns of use are well established, and they clearly illustrate how linear programming is presently being employed in the oil industry. Many refineries calculate optimal operating schedules on a weekly or a monthly basis. A number of oil companies apply a linear programming analysis whenever they are considering a major agreement for the exchange of crude or other products. Increasingly, oil firms are periodically analyzing their distribution patterns to discover transportation cost savings and profit potentials for new or expanded markets. The leading firms are using linear optimization models to test different strategies for long-term growth (e.g., five years). The world's dozen largest oil companies typically employ, at least, 25 to 35 people - and often two or three times that many - whose prime responsibility is to apply linear programming to the analysis of important decisions.

The worldwide dependence on linear programming by the petroleum industries suggests that this methodology could be used as an intelligence assessment tool. The objective of one of the projects in ORD's Energy Resources and Production Program was to determine the feasibility of assessing refinery production in denied areas. A large refinery in the Soviet Union was selected by the project team as the target. Various data sets such as capacities, utilities, crude types, and petrochemical facilities, were developed by the Agency's analysts. A linear programming model of the refinery

was developed by a contractor, and various production scenarios were developed to quantify the possible production streams. The successful demonstration of this phase of the project has led to a current follow-up effort designed to develop the capability and skills of Agency analysts to model intelligence problems dealing with refinery production on a local, nationwide, regional, or world-wide basis.

Appendix A provides a simple example which illustrates the application of linear programming for refinery modelling. It demonstrates how the flow of fluid through various units of a refinery generates the coefficients, variables, and equations of a linear program. For those readers who are inclined to delve into the mathematical aspects of linear programming, Appendix B provides an algebraic and geometric interpretation of the simplex algorithm.

IV. Data Generation and Management

Until this point, the emphasis of this paper has been on the applications of mathematical programming and, in particular, linear programming. However, many users of linear programming have found that their resources are now primarily committed, not to solving models, but to managing and controlling the volumes of data associated with the models. Recent developments in algorithmic and computer capabilities have enabled the practitioners to solve even larger problems with reductions in solution times. The problems that arise with these large models deal with the management of the data. Therefore, efficient systems are required for generating, manipulating, and analyzing large sets of data.

Systems for the generation and management of data have been and can be developed to serve the user's needs. The Defense and Intelligence communities have developed a system for managing the 100,000 data points that are collected over the Soviet Union each day by meteorological satellites. This data set is transformed into more useful information (temperatures, wind directions, cloud cover, etc.) for its many users.

The ORD funded linear programming model of the Soviet refinery consisted of 303 equations and 761 variables. A problem of this size is trivial to solve with the current software systems if the data are already generated and formatted. As small as this problem was, it potentially required over 230,000 nonzero technical coefficients with five decimal accuracy to be generated if the problem had a 100% dense matrix. Although the model formulation did not have anywhere near this density, there were, approximately, 10,000

coefficients which still had to be generated. This is not a trivial task. Our contractors estimate that an experienced modeler and a data clerk would have to spend four to six months generating this data by hand. However, utilizing their propriety systems, it took four days to generate the data and run a set of models. It is not uncommon for refinery modelers to develop and run models with 5,000 equations and 10,000 variables with the aid of their "complete" systems.

Probably one of the largest and most complex data management systems is utilized by the U.S. Department of Treasury. Two satistical data bases, Current Population Survey and Statistics of Income, are used extensively to analyze the effect of various policy changes such as welfare payments, social security benefits, and income tax rates. The two data bases do not contain the same information, but there is some overlap. Therefore, the files must be merged to be used for policy analysis. The information of this data management system is based on a network flow model. The size of the model is 50,000 nodes (equations) and 62.5 million links (variables). It takes three hours of central processing unit (CPU) and input-output time to make a run on a UNIVAC 1108 (1960 vintage) computer.

The preceding paragraphs illustrate the importance of developing database generators and management systems. ORD's goals have been not only to develop the appropriate methodologies, e.g., linear programming, but also to develop responsive data management systems. Only by these computer-based systems can the Intelligence Community efficiently handle and fully utilize the massive volume of data to which it has access.

APPENDIX A. LINEAR PROGRAMMING MODEL OF A REFINERY

The problem of optimizing refinery operations is extremely complex. Each batch of crude received at a refinery may possess different properties so that there are different proportions of yields of intermediate products. The refiner must determine which crude or mix of crudes should be run. When this decision is made, the output of the distillation unit is essentially determined. The next step is to determine what fractions of the distilled products should go directly to blending or cracking. Finally, at the blending stage, it must be determined how to blend the various streams and how much of the individual streams should be sold as final products.

Figures 1 and 2 illustrate the flow of fluids in a refinery with three main operations: distillation, cracking, and gasoline blending. Three different types of crude oil are processed by the refinery. Figure 2 shows the fractional breakdown of crude type no. 1 as it flows through each process; however, there is too much fuel, diesel, and stove oil produced and not enough of the other products to satisfy demand. Portions of these first three products are then sent to a thermal cracker to further reconfigure their molecular structures into the other desired lighter fractions. The last operation is the blending of gasoline. The proper mix of products for producing a specific gasoline is also shown in Figure 2.

The refinery is not only faced with determining an optimal routing of the various streams, but also with decisions about the

operating conditions of many of the process units. For this example, the operating conditions of each unit are assumed to be fixed. Thus, once the input to the unit is specified, then the various outputs are also determined.

The production goals of the refinery are reflected by the objective function. In many cases, it is not an easy task to decide what the objectives should be. The refinery may select criteria such as the following:

- o maximize profit
- o minimize cost
- o maximize gasoline production
- o minimize fuel oil production
- o maximize production of aviation fuels, fuel oils, and gasolines

Figure 3 illustrates the linear programming tableau (matrix) that is created by following the physical flow of products for this problem. The columns of the linear programming matrix are generated by the breakdown of the fluids by process and the process capacities. The breakdown by distillation for crude no. 1 appears under the variable X_4. The breakdown by cracking for fuel oil appears under the variable X_8. The requirements for a blended gasoline appear under variable X_{15}. The variables represent the process activities, unused capacities, and final products. The AVAILABLE column provides the capacity limits for processes and the material balance restriction constants for the fluids.

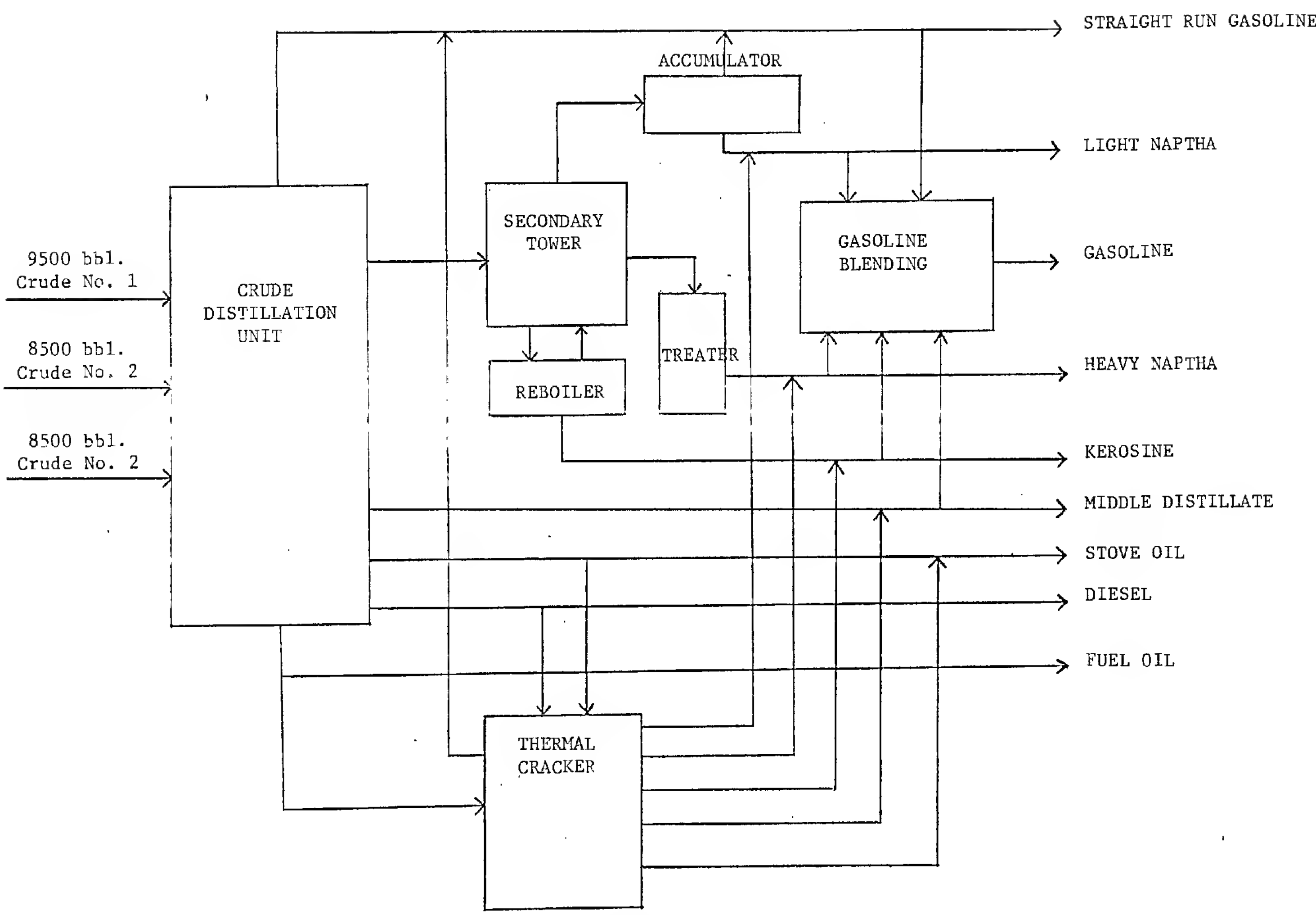
The rows represent the physical constraints on the refinery. The first row equation is: $X_1 + X_4 = 9500$. This equation can be

interpreted as follows:

(the number of barrels of crude no. 1 that are not used in a day) plus (the number of barrels of crude no. 1 that are distilled in a day) equals 9,500 barrels per day.

A similar interpretation can be applied to the process capacity constraints. The material balance constraints relate the input into any one process unit and the various output streams from the unit. The PROFIT row corresponds to the objective function which is to be maximized. The coefficients reflect the costs of operations (negative values) and the profit on the sale of products (positive values).

SIMPLIFIED REFINERY FLOW DIAGRAM



SECRET

FIGURE 1

SECRET

SEQUENCE FOR 1 BARREL OF CRUDE OIL NO. 1 THROUGH THE REFINERY

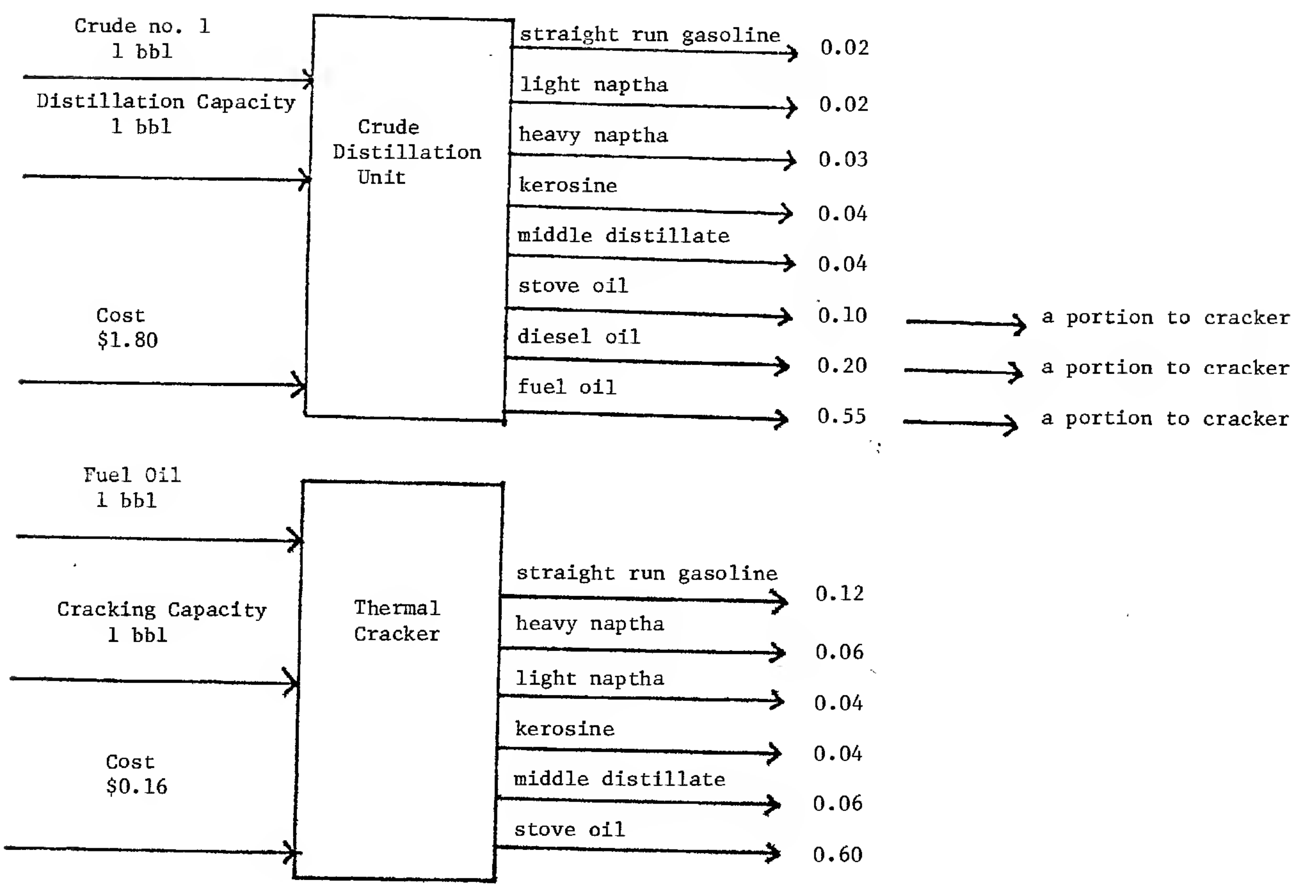


FIGURE 2

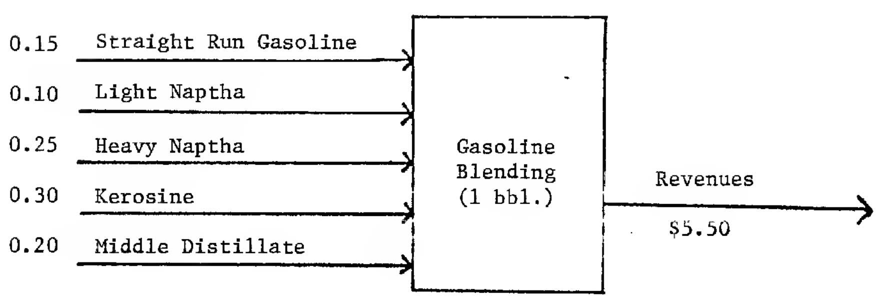


FIGURE 2 (Continued)

Approved For Release 2002/05/29 : CIA-RDP83T00573R000300220003-5

LINEAR PROGRAM OF REFINERY

ACTIVITIES / ITEMS	UNUSED AVAILABLE CRUDE OIL: CRUDE NO. 1	CRUDE NO. 2	CRUDE NO. 3	DISTILLATION: CRUDE NO. 1	CRUDE NO. 2	CRUDE NO. 3	UNUSED CAPACITY	CRACKING: FUEL OIL	DIESEL OIL	STOVE OIL	UNUSED CAPACITY	PRODUCT MARKETING: FUEL OIL	DIESEL OIL	STOVE OIL	GASOLINE BLENDING	MIDDLE DISTILLATE	KEROSINE	HEAVY NAPTHA	LIGHT NAPTHA	STRAIGHT RUN-GASOLINE	AVAILABLE bbl/day
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	
Crude no. 1	1			1																	=9500
Crude no. 2		1			1																=8500
Crude no. 3			1			1															=8000
Distillation capacity				1	1	1	1														=14,000
Material Balance Constraints For: -Fuel Oil				-.55	-.61	-.50		1				1									=0
-Diesel Oil				-.20	-.12	-.11			1				1								=0
-Stove Oil				-.10	-.07	-.14		-.6	-.2	1				1							=0
Cracking Capacity								1	1	1	1										=3500
Material Balance Constraints For: -Middle Distillate				-.04	-.06	-.05		-.06	-.41	-.30					.20	1					=0
-Kerosine				-.04	-.05	-.08		-.04	-.20	-.30					.30		1				=0
-Heavy Naptha				-.03	-.04	-.05		-.04	-.04	-.04					.25			1			=0
-Light Naptha				-.02	-.02	-.03		-.06	-.12	-.10					.10				1		=0
-Straight Run Gasoline				-.02	-.03	-.04		-.12	-.16	-.14					.15					1	=0
Profit				-1.8	-1.9	-2.0		-.16	-.21	-.21		1.8	4.0	4.2	5.5	4.0	4.1	4.2	4.3	3.3	= max.

FIGURE 3

Appendix B: Linear Programming and the Simplex Algorithm

The general problem of linear programming is the search for the optimum solution(s), either maximum or minimum, of a linear objective function of decision variables which must also satisfy a system of linear equations or inequalities. The formulation presented in the first section is the general representation of mathematical programming problems of which linear programming is but one class. A general algebraic formulation of a linear program can be represented as follows:

$$\text{MAXIMIZE} \quad \sum_{j=1}^{m} c_j x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{m} a_{ij}\, x_j \leq b_i \qquad i=1,\ldots,p$$

$$\sum_{j=1}^{m} a_{ij}\, x_j = b_i \qquad i=p+1,\ldots,m$$

$$x_j \geq 0 \qquad j=1,\ldots,q$$

$$\infty < X_j < \infty \ \underline{\text{(arbitrary)}} \qquad j=q+1,\ldots,n$$

where all of the a_{ij}, b_i, and c_j are numerical constants and x_j's are variables.

The only mathematical background requirements for a <u>complete</u> understanding of linear programming is a knowledge of linear algebra and matrix theory. The methods of classical optimization (differential calculus) are not appropriate for solving linear programs. These reasons will soon become apparent. However, no method has been found that will provide the optimal solution to a linear program in a single step. All of the techniques that have been developed over the years are iterative algorithms. The best known, most widely used, and most computationally efficient is the simplex algorithm.

Although a definitive and complete presentation of linear programming is beyond the scope of this paper, a brief and rudimentary explanation will be presented. This exposure will provide the geometrical concepts and the algebraic mechanics of the simplex algorithm. For simplicity, a linear program can be written in matrix notation as follows:

$$\begin{aligned} &\text{MAXIMIZE} && CX \\ &\text{subject to} && AX \leq b \\ & && X \geq 0 \end{aligned}$$

where A is an mXn matrix of the a_{ij}'s with m<n, X is an mX1 vector of the x_j's, C is an 1Xn vector of the c_j's, and b is an mX1 vector the b_j's.

A linear program may have only one of the following possible solutions:

(i) no feasible solution, that is, there are no values of all the x_j's that satisfy all of the constraints

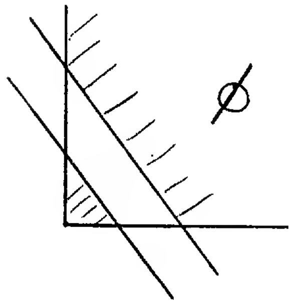
(ii) a unique optimal (feasible) solution

(iii) a nonunique optimal (feasible) solution; that is, convex combination of points give the same optimal solution

(iv) a feasible solution with an unbounded objective function; that is, by increasing certain variables indefinitely, the objective function can be increased indefinitely.

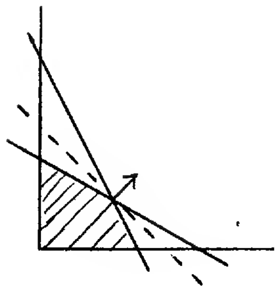
These cases can be represented and solved graphically in the two variable case. The first step is to find the solution space for the set of inequalities. If a finite optimum exists, then it will occur at one of the "corners" of the solution space. The maximum (minimum) point can be found by evaluating these corner

NO FEASIBLE SOLUTION



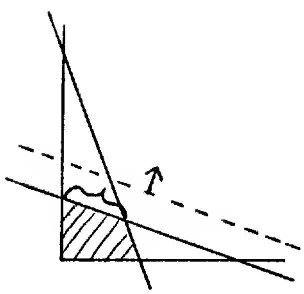
$$\text{MAXIMIZE} \quad 3X + 3X_2$$
$$\text{subject to} \quad 2X_1 + X_2 \leq 4$$
$$8X_1 + 5X_2 \geq 40$$
$$X_1 \geq 0, \; X_2 \geq 0$$

A UNIQUE OPTIMAL SOLUTION



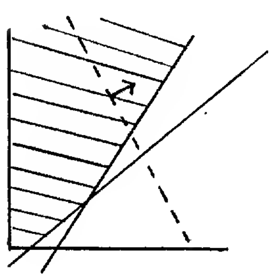
$$\text{MAXIMIZE} \quad X_1 + X_2$$
$$\text{subject to} \quad 10X_1 + 5X_2 \leq 50$$
$$4X_1 + 8X_2 \leq 32$$
$$X_1 \geq 0, \; X_2 \geq 0$$

A NONUNIQUE OPTIMAL SOLUTION



$$\text{MAXIMIZE} \quad 4X_1 + 14X_2$$
$$\text{subject to} \quad 2X_1 + 7X_2 \leq 21$$
$$7X_1 + 2X_2 \leq 21$$
$$X_1 \geq 0, \; X_2 \geq 0$$

UNBOUNDED OPTIMAL SOLUTION



$$\text{MAXIMIZE} \quad 2X_1 + X_2$$
$$\text{subject to} \quad X_1 - X_2 \leq 10$$
$$2X_1 - X_2 \leq 40$$
$$X_1 \geq 0, \; X_2 \geq 0$$

points in the objectives function or by parametrically graphing the objective function and locating the "furthest" corner point.

Some special definitions for matrices and vectors used in linear programming are required at this stage. A basis matrix B is an mXm matrix whose column vectors are composed of the column vectors from the mXn matrix A, i.e., B is a regular submatrix of order m of A. The m variables associated with the columns of the basis B are called the basic variables, will be postive in value, and denoted by X_B. The other (n-m) variables not associated with the basic matrix B are called nonbasic variables, will be zero, and will be denoted by the vector X_R. (A basic solution X_B is called degenerate if any of its variables have a zero value).

For a given basis B, the solution to the system of the constraints is as follows:

$$BX_B = b \quad \text{given } AX = b$$

$$\text{yields} \quad X_B = B^{-1}b \quad \text{as a solution}$$

$$X_R = 0$$

The method of calculating all possible basic solutions to find the solution of a linear program becomes infeasible when m or n exceed two or three. The total number of bases for a system of m equations and n variables is given by the combinatorial function:

$$C(n,m) = \frac{n!}{m!\,(n-m)!}$$

The calculation of all basic solutions for a problem of ten equations and twenty unknowns would require finding the solution to all possible ten by ten systems. There are 184,756 of these

systems. In fact, a large number of the systems would have no solutions or infeasible solutions. Even if this approach were programmed on a high speed computer, it still would require an enormous amount of computer time.

In contrast, a problem of this size can be solved by hand, albeit with some inconvenience, by using an iterative procedure such as the simplex algorithm. This problem becomes trivial for the large scale commerical linear programming packages and does not even pose problems for a small routine which an analyst could develop in most computer programming languages. In addition, the simplex algorithm would determine if the problem had an unbounded solution whereas the process of examining all possible basic solutions would not reveal such a fact.

The simplex method is an algebraic iterative procedure which will solve any linear programming problem exactly, disregarding roundoff errors, in a finite number of steps, or give an indication that there is an unbounded solution. Geometrically, the set of feasible solutions of the constraints represents a closed convex polyhedral set in the nonnegative orthant of Euclidean n-space. The vertices of this polyhedral set are determined by the set of all basic feasible solution vectors X_B. These vertices will be called extreme points. A unique optimal solution of a linear program will therefore be found at an extreme point. There is only a finite number of (feasible) extreme points. Assuming that the simplex algorithm begins at an extreme point (a basic feasible solution), the algorithm will move from an extreme point to an adjacent extreme point at each iteration. Of all

the possible adjacent extreme points, the one chosen is that which gives the greatest increase in the objective function (if the problem is being maximized). At each extreme point, the simplex algorithm indicates whether the extreme point is optimal, and if it is not, then which extreme point is the next choice. If the algorithm comes to an extreme point which has an edge leading to infinity and if the objective function can be increased indefinitely by moving along the edge, then the algorithm will indicate that fact.

The algebraic mechanics of the simplex algorithm can be represented efficiently in matrix and vector notation. By adding a nonnegative slack variable to each inequality constraint, the previously general linear program can be written as follows:

$$\begin{aligned} &\text{MAXIMIZE} \quad CX \\ &\text{Subject to } AX + IS = b \\ &X \geq 0 \\ &S \geq 0 \end{aligned}$$

where S is an mX1 vector of nonnegative variables and I is an mXm identity matrix. The conversion of this problem to a tabular vector representation upon which the simplex algorithm would be applied would have the following form:

$$\left[\begin{array}{c|cccccccc} 0 & -C_1 & -C_2 & \ldots -C_m & 0 & 0 & \ldots . 0 \\ b & A_1 & A_2 & \ldots A_m & I_1 & I_2 & \ldots I_m \end{array}\right]$$

where $\begin{bmatrix} 0 \\ I_j \end{bmatrix}$ is a column vector with a 1 in row j and zeros elsewhere and the A_j's are the vectors of the A matrix. This can be presented by a further condensation as:

$$\left[\begin{array}{c|cc} 0 & -C & \vec{0} \\ b & A & I \end{array}\right]$$

where $\vec{0}$ is an 1Xm vector of zeros. Let C_B denote the components of the vector C that are associated with the columns of the basic matrix B. The tableau representing the current iteration of the algorithm can be found by multiplying by the following basis inverse matrix:

$$\begin{bmatrix} 1 & -C_B \\ \vec{0} & B \end{bmatrix}^{-1} = \begin{bmatrix} 1 & C_B B^{-1} \\ 0 & B^{-1} \end{bmatrix}$$

The current tableau of a basic feasible solution is found by:

$$\begin{bmatrix} 1 & C_B B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 0 & -C & 0 \\ b & A & I \end{bmatrix}$$

$$\begin{bmatrix} C_B B^{-1} b & C_B B^{-1} A - C & C_B B^{-1} \\ B^{-1} b & B^{-1} A & B^{-1} \end{bmatrix}$$

The value of the objective function is given by $C_B B^{1} b$. The basic variables are found, as before, $X_B = B^{-1} b$, and $X_R = 0$.

Since the algorithm moves from an extreme point to an adjacent extreme point, then the first step is to select the variable that is to enter the basis and which gives the greatest increase in the objective function. This is accomplished by selecting the nonbasic variable associated with the most negative entry in the set $[C_B B^{-1} A - C \quad C_B B^{-1}]$ given in the simplex tableau. The second step is to select the basic variable which

will leave the basis. If the k-th nonbasic variable is selected to enter the basis, then the positive elements of that column are divided into the corresponding elements of the column vector $B^{-1}b$. The basic variable associated with the minimum positive ratio of the division is the variable selected to leave the basis. Assume this occurs in row k. Assume the column vector of the j-th variable is represented as follows:

$$\begin{bmatrix} C_{oj} \\ r_{1j} \\ \vdots \\ r_{kj} \\ \vdots \\ r_{mj} \end{bmatrix}$$

The ensuring change-of-basis calculation (new simplex tableau) can be found by multiplying the tableau by the (m+1) X (m+1) elementary transformation matrix E where the ratios occur in the (k+1)st column:

$$E = \begin{bmatrix} 1 & 0 & \cdots & -C_{oj}/r_{kj} & \cdots & 0 \\ 0 & 1 & \cdots & -r_{1j}/r_{kj} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/r_{kj} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & -r_{mj}/r_{kj} & \cdots & 1 \end{bmatrix}$$

The solution will be optimal when $C_B B^{-1} A - C \geq 0$ and $C_B B^{-1} \geq 0$ in the tableau.

BIBLIOGRAPHY

Adams, F. and J. Griffin, "An Economic-Linear Programming Model of the U.S. Petroleum Refining Industry," Journal of the American Statistical Association, Vol. 67, No. 339, 1972, pp. 542-551.

Charnes, A. and W. W. Cooper, Management Models and Industrial Applications of Linear Programming, 2 vols., New York: John Wiley and Sons, Inc., 1961.

Charnes, A., W. W. Cooper, and B. Mellon, "A Model For Programming And Sensitivity Analysis In An Integrated Oil Company," Econometrica, Vol. 22, No. 2, 1954, pp. 193-217.

Dantzig, G., Linear Programming and Extensions, Princeton, New Jersey: Princeton University Press, 1963.

Glover, F., J. Hultz, and D. Klingman, "Improved Computer-Based Planning Techniques, Part I," Interfaces, Vol. 8, No. 4, 1978, pp. 16-25.

Glover, F. and D. Klingman, "Network Application in Industry and Government," AIIE Transactions, Vol. 9, No. 4, 1977, pp. 363-376.

Glover, F., G. Jones, D. Karney, D. Klingman, and J. Mote, "An Integrated Production, Distribution, and Inventory Planning System," Interfaces, Vol. 9, No. 5, 1979, pp. 21-35.

Hadley, G., Linear Programming, Reading, Massachusetts: Addison-Wesley Publishing Company, 1962.

Manne, A., "A Linear Programming Model Of The U.S. Petroleum Refining Industry," Econometrica, Vol. 26, No. 1, 1958, pp. 67-106.

Thomas, G. and J. DaCosta, "A Sample Survey of Corporate Operations Research," Interfaces, Vol. 9, No. 4, 1979, pp. 102-111.

Wagner, H., Principles of Operations Research With Applications To Managerial Decisions, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1969.

Zierer, T., W. Mitchell, and T. White, "Practical Applications of Linear Programming To SHELL's Distribution Problems, Interfaces, Vol. 6, No. 4, 1976, pp. 13-26.

25X1